Warm Up # 6-1 # 4 from p. 346

4 A clothing store recorded the length of time customers were in the store and the amount of money they spent.

Time (min)	8	18	5	10	17	11	2	13	18	4	11	20	23	22	17
Money (€)	40	78	0	46	72	86	0	59	33	0	0	122	90	137	93

a Draw a scatter diagram of the data. **b** Calculate the mean point.

c Plot the mean point on your diagram and draw a line of best fit through the mean point.

d Describe the relationship between *time in the store* and *money spent*.

e Estimate the amount of money spent by a person who is in the store for 15 minutes. Comment on the reliability of your estimation.

HW Questions: p. 341

EXERCISE 11E.2

1 This contingency table shows the responses of a randomly chosen sample of adults regarding the person's weight and whether they have *diabetes*.

At a 5% significance level, the critical value of χ^2 is 5.99.

Test at a 5% level whether there is a link between *weight* and suffering *diabetes*.

	Weight		
	light	*medium*	*heavy*
Diabetic	11	19	26
Non-diabetic	79	68	69

2 The table opposite shows the way in which a random sample of people intend to vote in the next election.

	Age of voter		
	18 to 35	36 to 59	60+
Party A	85	95	131
Party B	168	197	173

a For a 10% significance level, what is the critical value of χ^2?

b Test at a 10% level whether there is any association between the *age of a voter* and the *party they wish to vote for*.

3 Noah wanted to find out whether there is a relationship between a person's *gender* and their *favourite season*. He sampled 100 people, and obtained the results alongside.

	Favourite season			
	Summer	*Autumn*	*Winter*	*Spring*
Male	8	11	6	20
Female	12	17	10	16

At a 1% significance level, the critical value for this test is 11.34.

Test, at a 1% level, whether the variables *gender* and *favourite season* are independent.

4 The guests staying at a hotel are asked to provide their *reason for travelling*, and to *rate* the hotel on a scale from Poor to Excellent. The results are shown below.

		Rating			
		Poor	*Fair*	*Good*	*Excellent*
Reason for travelling	*Business*	27	25	20	8
	Holiday	9	17	24	30

a Show that, at a 5% significance level, the variables *reason for travelling* and *rating* are dependent.

b By examining the contingency table, describe how a guest's *rating* is affected by their *reason for travelling*.

5 The hair and eye colours of 150 randomly selected individuals are shown in the table below.

		Hair colour			
		Blond	*Black*	*Brunette*	*Red*
Eye colour	*Blue*	14	10	21	5
	Brown	11	32	20	12
	Green	5	2	14	4

At a 5% significance level, the critical value for χ^2 is 12.59.

Test, at a 5% level, whether there is an association between *hair colour* and *eye colour*.

6 Hockey player Julie wondered whether the position you played affected your likelihood of being injured. She asked a random sample of hockey players what position they played, and what injuries they had sustained in the last year.

		Position			
		Forward	*Midfielder*	*Defender*	*Goalkeeper*
Injury type	*No injury*	23	18	24	7
	Mild injury	14	34	23	11
	Serious injury	10	16	13	7

Test, at a 10% significance level, whether the variables *position* and *injury type* are independent.

Review:

Chi Square Test of Independence.

X^2 is a statistic that measures the difference between observed values and expected values in a contingency table.

If the calculated Chi Square value is big enough, we can establish a:

link between two variables

association between two variables

relationship between two variables

Independent ⟷ Not Independent

~~Dependent~~

The test is for Categorical variables only.

- Set up a contingency table for two categorical variables.
- Assume independence to start
- Calculate expected values

We only use the term "correlation" with with numerical data.

The cutoff, or critical Chi-Square Value, is either given to you or found in a resource table.

This value will tell you whether to accept or reject the assumed independence between the two variables.

THE FORMAL TEST FOR INDEPENDENCE

Step 1: State H_0 called the **null hypothesis**. This is a statement that the two variables being considered are independent.
State H_1 called the **alternative hypothesis**. This is a statement that the two variables being considered are not independent.

Step 2: State the **rejection inequality** $\chi^2_{calc} > k$ where k is the **critical value** of χ^2.

Step 3: Construct the expected frequency table.

Step 4: Use technology to find χ^2_{calc}.

Step 5: We either reject H_0 or do not reject H_0, depending on the result of the rejection inequality.

Step 6: We could also use a ***p*-value** to help us with our decision making.
For example, at a 5% significance level: If $p < 0.05$, we reject H_0.
If $p > 0.05$, we do not reject H_0.

If the p-value is smaller than the significance level, then it is sufficiently unlikely that we would have obtained the observed results if the variables had been independent. We therefore conclude that the variables are not independent.

p-value: probability of getting observed values as far, or further, from the expected values (assuming independence)

Classwork: Read p. 342, Limitations of the Chi Square Test

Do 11E.3 p.343, # 1 & 2

11E.4 p.344 # 1

Note: This classwork will be carried over to next week.

HW: 11D p.333, # 5

Rev. Set 11A p.345, # 1 - 3, 7

HW Quiz Tuesday

pgs. 541, 59, 62, 336, 337, 341

Quiz Tuesday:

Normal Distribution

Correlation

X^2 Test of Independence